HAMILTON CIRCUITS WITH MANY COLOURS IN PROPERLY EDGE-COLOURED COMPLETE GRAPHS

LARS DØVLING ANDERSEN*

Abstract.

We prove that a properly edge-coloured complete graph $K_n$ has a Hamilton circuit with edges of at least $n - \sqrt{2n}$ distinct colours. This is proved with a method inspired by work on long partial transversals in latin squares. Another such method is employed in proving a similar result where the distinct colours all occur on edges not belonging to a given spanning set of edges of $K_n$.

1. Introduction.

The purpose of this paper is threefold. We consider the complete graph $K_n$ with a proper edge-colouring with any number of colours (as the edge-colouring is proper, the number of colours is at least $n - 1$ if $n$ is even, and at least $n$ if $n$ is odd), and we look for a Hamilton circuit of $K_n$ with as many different colours as possible occurring on its edges. Firstly, we want to give a lower bound on this number; in Section 3 it is proved that for all $n$ it is possible to get at least $n - \sqrt{2n}$ distinct colours occurring in a Hamilton circuit. Secondly, we wish to indicate how methods from the theory of long partial transversals in latin squares can be adapted to the present problem; Section 2 briefly introduces such transversals and mentions some results. Finally, with two applications in mind, we prove a result similar to the main theorem, but where some given edges are “forbidden”.

We use standard terminology and state only a few definitions. By a path system of a graph we mean a subgraph consisting of mutually vertex-disjoint paths; its length is the sum of the lengths of the paths, i.e. the number of edges. A subgraph of an edge-coloured $K_n$ is called totally multicoloured (TMC) if all its edges have distinct colours. If $K_n$ does not contain a TMC Hamilton circuit, then asking for the maximum number of distinct colours occurring in a Hamilton circuit is

* This research was carried out while the author held a Niels Bohr Fellowship from The Royal Danish Academy of Sciences and Letters.

Received December 12, 1985; in revised form September 20, 1987.
equivalent to asking for the maximum length of a TMC path system, and it is in
this form that we shall tackle the problem.

An edge-colouring of a graph with \( c \) colours is an assignment to each edge of
a colour from a set of \( c \) colours. We shall often use the set of "colours" \( \{1, 2, \ldots, c\} \).
An edge-colouring is proper if no pair of edges incident with the same vertex have
the same colours. We emphasize that this paper is concerned with proper edge-
colourings. If the condition that the edge-colouring be proper is dropped, we
enter the field of anti-Ramsey theory; for example, the problem of finding a TMC
path of maximum length in a \( K_n \) with a (general) edge-colouring is considered by
M. Simonovits and V. T. Sós in [21]. More anti-Ramsey theorems can be found
in [9]. Another possibility is to replace the properness condition by an upper
bound on the number of occurrences of each colour; this is sub-Ramsey theory.
G. Hahn and C. Thomassen [14] proved that there is a constant \( c \) such that the
following holds: If each colour of an edge-colouring of \( K_n \) occurs on at most \( cn^t \)
edges, then \( K_n \) has a TMC Hamilton circuit. Recently, V. Rödl and independently
P. Winkler have improved the bound to \( c \cdot \sqrt{n} \) (private communication with
V. Rödl and G. Hahn). More sub-Ramsey results can be found in [12] and [13].

2. Partial transversals in latin squares.

A partial transversal of a latin square is a set of cells in distinct rows, in distinct
columns and containing distinct entries. The length of a partial transversal is the
number of cells in it. A transversal of a latin square of side \( n \) is a partial transversal
of length \( n \).

A conjecture due to H. J. Ryser [19] and also to R. A. Brualdi (see [7], page
103) states that any latin square of odd side has a transversal, and that any latin
square of even side \( n \) has a partial transversal of length at least \( n - 1 \). As early as
1779 L. Euler [10] proved that a cyclic latin square of even side has no transversal,
and so the statement of the conjecture is the best that could be hoped for.
Later, E. Maillet [17] gave other examples of latin squares of even side with no
transversals.

Work towards a proof of the conjecture has mainly been in the form of lower
bounds on the length \( p \) of a longest partial transversal of a latin square of side \( n \). It
is trivial that \( p \geq \frac{n}{2} \) (any partial transversal of length less than \( \frac{n}{2} \) can be
extended). K. K. Koksm [15] proved that \( p \geq \frac{2}{3} n + \frac{1}{3} \) for all \( n \geq 3 \), and D. A.
Drake [8] improved this to \( p \geq \frac{1}{4} n \) for all \( n \geq 8 \) (he actually proved \( p \geq \min \{ \frac{3}{4} n, \}
\( n - 2 \} \) for all \( n \geq 1 \) ). Later, A. E. Brouwer, A. J. de Vries & R. M. A. Wieringa [6]
and independently D. E. Woolbright [22] showed that \( p \geq n - \sqrt{n} \) for all \( n \) (the
bound can be worked out to give \( p \geq n + \frac{1}{2} - \sqrt{n - \frac{3}{4}} \) for all \( n \geq 3 \) , [6]).

The latest improvement is a result by P. W. Shor [20] saying that a latin square
of side $n$ has a partial transversal of length $p \geq n - c(\log n)^2$, where $c$ is a constant close to 5.53. This is better than $n - \sqrt{n}$ for $n \geq 2 \cdot 10^6$.

If we now consider TMC path systems of edge-coloured complete graphs, we can see that they have certain similarities with partial transversals of Latin squares. In particular, both properties are hereditary: a subgraph of a TMC path system is itself a TMC path system, and a subset of a set of cells forming a partial transversal also form a partial transversal. Further, in both cases we must choose a structure for which certain symbols must all be distinct, in a larger structure which is defined by the requirement that some symbols must be different. So perhaps it is no surprise that methods from the theory of partial transversals often have close analogues in the theory of TMC path systems.

In this paper we shall present such analogues of the methods of Brouwer, de Vries and Wieringa (for the proof of our main result in Section 3) and of Drake (for the application in Section 4). Unfortunately, we have not been able to adapt the method of Shor to obtain further improvements. So let us stress that there are also huge differences between partial transversals and TMC path systems: most notably, the symmetry among all cells of a partial transversal does not exist among the edges of a path system (some are end-edges, some are not, for example).

3. Long totally multicoloured path systems.

Hahn [12] conjectured that for any (general) edge-colouring of $K_n$ in which each colour occurs on at most $k$ edges, where $n \geq 2k$ (and $n \not\equiv 4$), $K_n$ contains a TMC Hamilton path. This was disproved by M. Maamoun and H. Meyniel [16] for all $n$ of the form $n = 2^p$, and in fact their counterexample had a proper edge-colouring: Let the vertices of $K_{2^p}$ be the elements of the group $(\mathbb{Z}_2)^p$, and let the set of colours be the same elements, zero excepted; for all vertices $x$ and $y$ the edge joining them is given the colour $x - y$. Then, if $x_1x_2\ldots x_n$ were a TMC Hamilton path we would have

$$x_n - x_1 = (x_n - x_{n-1}) + (x_{n-1} - x_{n-2}) + \ldots + (x_2 - x_1) = \sum_{x \neq 0} x = 0$$

(where the sum is taken over all elements of $(\mathbb{Z}_2)^p$), a contradiction.

There are other counterexamples to Hahn’s conjecture. In fact, the unique proper edge-colouring of $K_6$ with 5 colours does not even contain a TMC path system consisting of two disjoint 2-paths.

It may be true, however, that a properly edge-coloured $K_n$ always has a TMC path system of length $n - 2$, or even of length $n - 1$ if $n$ is odd (see Section 5). It is trivial that it contains a TMC path of length at least $\frac{n}{2}$; a rather simple counting argument gives a bound of $\frac{3}{2}(n - 2)$. Methods inspired by Drake’s paper [8] give
\(3n - 5 \frac{4}{4}\) (we return to these methods in the next section), but the best result that we have is based upon the methods of Brouwer, de Vries and Wieringa [6]:

**Theorem 1.** Let \(K_n\) have a proper edge-colouring with any number of colours. Then it contains a totally multicoloured path system with at least \(n - \sqrt{2n}\) edges.

**Proof.** Let the colours be \(\{1, 2, \ldots, c\}\), \(c \geq n - 1\), and let \(P\) be a TMC path system of maximum length \(p\). Let the edges of \(P\) be denoted by \(e_1, \ldots, e_p\), named consecutively along each path (component) belonging to \(P\); with the same way of traversing each path, let \(V_i\) be the last end-vertex of \(e_i\), \(1 \leq i \leq p\). We may assume that \(e_i\) has colour \(i\), \(1 \leq i \leq p\). Let \(W = \{V_1, \ldots, V_{p}\}\), and let the set of remaining vertices be \(U = \{V_{p+1}, \ldots, V_n\}\). Note that \(W \neq V(P)\). For \(1 \leq i \leq n, 1 \leq j \leq n, i \neq j,\) let \(c(V_i, V_j)\) be the colour of the edge joining \(V_i\) and \(V_j\). Figure 1 illustrates the notation. The labelling of the vertices of \(U\) is arbitrary.

![Diagram](image)

Figure 1.

Define the following sets (they may be thought of as corresponding to sets of vertices of \(K_n\)):

\[
A_0 = \emptyset, \\
A_j = \{i \in \{1, \ldots, p\} \cup \{p + j + 1, \ldots, n\} \mid c(V_i, V_{p+j}) \notin \{1, \ldots, p\} \setminus A_{j-1}\}
\]

for \(1 \leq j \leq n - p\).

Clearly \(A_j = \{i \in \{1, \ldots, p\} \cup \{p + j + 1, \ldots, n\} \mid c(V_i, V_{p+j}) \in \{p + 1, \ldots, c\} \cup A_{j-1}\}\).

We claim:

\((\star)\) \quad \{p + 1, \ldots, n\} \cap A_j = \emptyset \text{ for all } j \in \{0, \ldots, n-p\}.

**Proof of \((\star)\)** Assume that \((\star)\) is false, and let \(j_0\) be the smallest \(j\) for which it fails; obviously \(j_0 \geq 1\). Let \(q_0 \in \{p + 1, \ldots, n\} \cap A_{j_0}\). By the definition of \(A_{j_0}\), \(q_0 \geq p + j_0 + 1\). Then \(P \cup \{V_{q_0}, V_{p+j_0}\}\) is also a path system, and so by the maximality of \(P\), \(c(V_{q_0}, V_{p+j_0})\) belongs to \(\{1, \ldots, p\}\). As \(q_0 \in A_{j_0}\) this means that \(c(V_{q_0}, V_{p+j_0}) \in A_{j_0-1}\). Put \(q_1 = c(V_{q_0}, V_{p+j_0})\), and let \(j_1\) be the smallest number so that \(q_1 \in A_{j_1}\). Put \(P_1 = (P \cup \{V_{q_0}, V_{p+j_0}\}) \setminus \{e_{q_1}\}\). Then \(P_1 \cup \{V_{q_1}, V_{p+j_1}\}\) is again
a path system with the same colours as $P$, and as $q_i \in A_j$, the maximality of $P_1$ implies that $q_2 = c(V_{q_i}, V_{p+j_i}) \in A_j - 1$. Continuing like this we reach, in a finite number of steps, a contradiction to the fact that $A_0 = \emptyset$.

For all $j$, $1 \leq j \leq n - p$, we have by (*) and the definition of $A_j$:

$$|A_j| \geq p + (n - p - j) - (p - |A_{j-1}|) = n - j - p + |A_{j-1}|$$

so we get

$$|A_{n-p}| = \sum_{j=1}^{n-p} (|A_j| - |A_{j-1}|)$$

$$\geq \sum_{j=1}^{n-p} (n - j - p)$$

$$= (n - p)^2 - \frac{1}{2}(n - p)(n - p + 1)$$

$$= \frac{1}{2}(n - p)(n - p - 1).$$

As $|A_{n-p}| \leq p$ by definition, we get $\frac{1}{2}(n - p)(n - p - 1) \leq p$ which implies

$$p \geq n + \frac{1}{2} - \sqrt{2n + \frac{1}{4}}.$$  

This completes the proof of Theorem 1.

The proof of Theorem 1 actually gives $p \geq n + \frac{1}{2} - \sqrt{2n + \frac{1}{4}}$, and combining this with the bound $\frac{3n - 5}{4}$ derived by methods inspired by Drake [8] (the methods are indicated in the next section), we obtain as our best possible result:

**Corollary 2.** Let $K_n$ have a proper edge-colouring with any number of colours. Then $t$ contains a totally multicoloured path system of length at least

$$\frac{3n - 5}{4} \text{ if } n \leq 15,$$

$$n + \frac{1}{2} - \sqrt{2n + \frac{1}{4}} \text{ if } n \geq 15.$$  

4. TMC path systems avoiding a spanning edge-set and containing a given TMC path system.

In two forthcoming papers [3] and [4], coauthored with A. J. W. Hilton and E. Mendelsohn respectively, we prove results about extending proper edge-colourings of some edges of $K_n$ to proper edge-colourings of all of $K_n$. The statements may be found in [1]. An important tool for both papers is a lemma ascertaining the existence of a long TMC path system in a properly edge-coloured complete graph, with the extra condition that all edges of a fixed spanning edge-set must be avoided by the TMC path system (these edges are the precoloured edges).
A similar result is needed with yet a further condition, namely that one or two given edges are mandatory for the TMC path system. In this section we sketch the proof of a lemma applicable to all cases needed in [3] and [4]. As was the case with Theorem 1, the proof is based on methods previously used for long partial transversals in latin squares; again, the methods due to Brouwer, de Vries and Wieringa, and also to Woolbright, give the best results asymptotically. It turns out, however, that the results obtainable by Drake's methods are more suited for the applications: they are strong enough for large \( n \), and they work for some small \( n \) where the asymptotically best results fail. So we sketch a proof based on Drake's ideas, thus presenting the adaptation of these for TMC path systems. Afterwards we state what can be obtained by other methods.

It can be noted that [1] and [2] contain a result on long partial transversals of (possibly incomplete) latin squares avoiding certain cells of the latin squares.

**Lemma 3.** Let \( K_n \) have a proper edge-colouring with any number of colours \( (n \geq 2) \), let \( F \) be a spanning edge-set of \( K_n \) and let \( M \) be the edges of some totally multicoloured path system of \( K_n \), where \( F \cap M = \emptyset \). Put

\[
\varepsilon(n, F, M) = \frac{1}{8}(\sqrt{n^2 - (22 - 4|M|)n} + 96|F| + 4(|M|^2 + |M|)) + 9 - \frac{n}{n + 2|M|} + 5).
\]

Then \( K_n \) contains a totally multicoloured path system, avoiding all edges of \( F \) and including all edges of \( M \), of length at least

\[
\min \{n - 5, \frac{3n}{4} - \varepsilon(n, F, M)\}.
\]

**Proof.** For the sake of brevity, let us call a TMC path system, avoiding all edges of \( F \) and including all edges of \( M \), an \((F, M)\)-system. Put \(|F| = f\) and \(|M| = m\). The edges of \( F \) are called forbidden.

Trivially, \( K_n \) contains an \((F, M)\)-system. Let \( P \) be such a system with the maximum number \( p \) of edges. If \( p \geq n - 5 \) we are finished, so we now assume that \( p \leq n - 6 \). We must prove that \( p \geq \frac{3n}{4} - \varepsilon(n, F, M) \).

Let the notation be as in the proof of Theorem 1, indicated in Figure 1. We call the colours 1, \ldots, \( p \) small colours and the colours \( p + 1, \ldots, c \) large colours. By the maximality of \( P \), no non-forbidden edge joining two vertices of \( U \) can have large colour.

Each connected component of \( P \) has one vertex in \( U \). We call this the \( U \)-vertex of the path and of its vertices and edges.

For each edge \( e_i \) of \( P \), let \( g(e_i) \) denote the number of edges of large colour which are not in \( F \) and which join an end-vertex of \( e_i \) to a vertex of \( U \), excluding the edge joining \( V_{i-1} \) to its \( U \)-vertex; note that this exception can only apply if \( V_{i-1} \) is an end-vertex of \( e_i \) (i.e., if \( e_i \) is not the first edge of its component).
We now partition the edges of $P$ into 4 mutually disjoint sets:

\[
M, \\
L = \{ e_i \in E(P) \setminus M \mid g(e_i) = 0 \}, \\
K = \{ e_i \in E(P) \setminus M \mid 1 \leq g(e_i) \leq 4 \}, \\
J = \{ e_i \in E(P) \setminus M \mid g(e_i) \geq 5 \}.
\]

We put $|L| = l$, $|K| = k$ and $|J| = j$.

Then the following statements can be proved:

1. \( e_i \in M \) \( \Rightarrow \) \( g(e_i) \leq 2(n - p) - 1 \).
2. \( e_i \in J \) \( \Rightarrow \) \( g(e_i) \leq n - p \).
3. \( e_i \in J \) \( \Rightarrow \) the colour $i$ does not occur on any non-forbidden edge joining two vertices of $U$.
4. \( e_i \in K \) \( \Rightarrow \) the colour $i$ occurs on at most one non-forbidden edge joining two vertices of $U$.

We also need an observation regarding the set $F$. Let $f_1$ be the number of forbidden edges joining a vertex of $W$ to a vertex of $U$, and let $f_2$ be the number of forbidden edges joining two vertices from $U$. From the fact that $F$ is spanning, (5) can be proved:

\[
f_1 + 2f_2 \leq 2f - p.
\]

Let $h$ be the number of non-forbidden edges of large colour joining a vertex of $W$ to a vertex of $U$. We shall compare $2h$ to the sum $\sum_{i=1}^{n} g(e_i)$. A counting argument shows:

\[
2h \leq \sum_{i=1}^{n} g(e_i) + p - 1
\]

Applying (1), (2) and the definition of $K$ we obtain:

\[
2h \leq (n - p)j + 4k + (2(n - p) - 1)m + p - 1.
\]

In addition to this upper bound on $h$, we can get the following lower bound:

\[
h \geq (n - p)(n - p - 1) - 2f_2 - f_1.
\]

With (7), this gives

\[
(n - p)j + 4k \geq 2(n - p)(n - p - 1) - (2(n - p - 1)m - p + 1 - 4f_2 - 2f_1).
\]

No large colour occurs on a non-forbidden edge joining two vertices of $U$, and by (3) no colour from $\{ i \mid e_i \in J \}$ is on such an edge. By (4), any colour from
\{i| e_i \in K\} occurs on at most one of these edges, and as the edge-colouring is proper each colour of \(\{i| e_i \in L \cup M\}\) occurs on at most \(\frac{n - p}{2}\) such edges. Therefore

\[(10)\quad k + \frac{n - p}{2}(l + m) \geq (n - p)(n - p - 1) - f_2.\]

Multiplying (10) by 2 and adding (9) gives

\[(11)\quad (n - p)(j + l + m) + 6k \geq 3(n - p)(n - p - 1) - 2(n - p)m + m - p + 1 - 6f_2 - 2f_1.\]

Applying \(n - p \geq 6\) and \(2f_1 + 6f_2 \leq 6f - 3p\) (from (5)), we get

\[p = j + k + l + m \geq 3(n - p - 1) - 2m - 2 - \frac{6f - 2n - m - 1}{n - p}\]

and

\[4p \geq 3n - 5 - 2m - \frac{6f - 2n - m - 1}{n - p}.\]

Looking at this as a second degree inequality in \(p\) (or \(n - p\)) it can be seen to imply the statement of the lemma.

The rather detailed expression for the bound in Lemma 3 is necessary for the applications. Putting \(M = \emptyset\) gives a result purely on TMC path systems avoiding certain edges. It is not possible to put \(F = \emptyset\), as \(F\) is required to be spanning.

Putting \(|F| = \frac{n}{2}\) and \(M = 0\) gives a bound of approximately \(\frac{3n - 8}{4}\), close to the corresponding part of Corollary 2.

As mentioned before, the method of the proof of Theorem 1 can also be extended to the case of forbidden and mandatory edges, and they give asymptotically better results. Lemma 3, however, is the more useful for the applications in [3] and [4]. Including both results, we can state the following theorem.

**Theorem 4.** Let \(K_n, F, M\) and \(\varepsilon(n, F, M)\) be as in Lemma 3. Then \(K_n\) contains a totally multicoloured path system, avoiding all edges of \(F\) and including all edges of \(M\), of length at least

\[\max\left\{n - |M| - \frac{1}{2} - \sqrt{4|F| + |M|^2 - |M| + \frac{1}{4}}, \min\left\{n - 5, \frac{3n}{4} - \varepsilon(n, F, M)\right\}\right\}.\]

If we put \(|M| = 0\) and \(|F| = \frac{n}{2}\) in the first term of the maximum clause, we obtain a bound one below the second bound of Corollary 2. That is, if \(n\) is even the
bound obtained by the method of the proof of Theorem 1 is just weakened by one by the additional requirement that a given 1-factor of $K_n$ must be avoided by the TMC path system.

We finally note that the Drake method can be extended so as not to require $n - p \geq 6$. This was done to obtain the Drake part of Corollary 2.

5. Concluding remarks

As mentioned in Section 3, there is room for improvement on the main result of this paper, Theorem 1. In fact, we believe in the following conjecture.

CONJECTURE. A properly edge-coloured $K_n$ has a TMC path of length at least $n - 2$.

This has been verified by hand for $n \leq 8$. Further evidence in favour of the conjecture is that it holds for the standard (cyclic) 1-factorization of $K_{2n}$ called $GK_{2n}$ (see [18]); for $n \geq 2$, it obviously has a TMC circuit of length $n - 1$. The conjecture also holds for the edge-colouring of $K_{2n}$ described in Section 3, which does not have a TMC path of length $n - 1$. It follows from the fact that the group $(Z_2)^n$ is $R$-sequenceable (which was proved by R. J. Friedlander, B. Gordon and M. D. Miller in [11]), that it also contains a TMC circuit of length $n - 1$.

Maybe new methods developed for finding long partial transversals in latin squares can be used on the above problem. And, perhaps more intriguing, maybe new methods developed for finding a long TMC path system in a properly edge-coloured complete graph can be used on the long standing problem of finding long transversals in latin squares.

We finally mention two results slightly related to the topics of this paper:

In [5], A. E. Brouwer used methods similar to those of Section 3 for finding a large partial parallel class in a Steiner triple system. In [23], D. E. Woolbright showed that in any 1-factorization of $K_{2n}$ there is a 1-factor containing edges from at least $n - 1$ distinct 1-factors of the 1-factorization. As Woolbright's methods apply to any proper edge-colouring, we can actually write the following corollary:

COROLLARY 5. Let $K_{2n}$ have a proper edge-colouring with any numbers of colours. Then it contains a set of $n - 1$ independent edges of distinct colours.

Clearly, we also have

COROLLARY 6. Let $K_{2n+1}$ have a proper edge-colouring with any number of colours. Then it contains a set of $n - 1$ independent edges of distinct colours.

Recently, Woolbright has improved his result to state that in any 1-factorization of $K_{2n}$ there is a 1-factor all of whose edges belong to different 1-factors of the 1-factorization (private communication with C. C. Lindner).
ACKNOWLEDGEMENT. A. J. W. Hilton, A. D. Keedwell and P. Landrock are thanked for very helpful discussions.

REFERENCES

4. L. D. Andersen & E. Mendelsohn, 1-factorizations of $K_{2m}$ with given edges in distinct 1-factors, in preparation.
10. L. Euler, Recherches sur une nouvelle espèce de quarrés magiques, in Leonard Eulier Opera Omnia, Série 1, 7 (1920), 291–392.
22. D. E. Woolbright, An $n \times n$ latin square has a transversal with at least $n - \sqrt{n}$ distinct symbols, J. Combin. Theory (A) 24 (1978), 235–237.

INSTITUTE OF ELECTRONIC SYSTEMS
AALBORG UNIVERSITY CENTRE
STRANDVEJEN 19
DK-9000 AALBORG
DENMARK